Randomized Methods in Computational Geometry

(Marco Pellegrini)

Lecture 6: Tail Estimates and Configuration Spaces

1

Elements of Probability Theory

We give now a self-contained treatment of some results of probability theory that we will need.

Given a discrete random Variable X taking values in

a discrete domain D $\subseteq$ Z, E[X] = be its expected value:

$$E[X] = \sum_{x \in D} x \cdot \Pr[X = x]$$

We are interested in bounding the probability that X

has a large deviation from E[X].

2

For X a non-negative r.v. Markov's inequality states:

$$\Pr[X \geq t] \leq \frac{E[X]}{t}$$

Proof: follows directly from the definition of E[X].

Let σ^2 = E[X²] - E[X]², σ is the standard deviation. Chebychev inequality states:

$$\Pr[|X - E[X]| \geq t\sigma] \leq \frac{1}{t^2}$$

Proof: follows from applying Markov's inequality to the random variable (X - E[X])².

3

Chernoff's technique

- It is often possible to obtain tighter bounds when the r.v. X can be expressed as sum of other r.v. X = $\Sigma_i X_i$, where the X_i are <u>independent</u>.
- The basic idea is to consider instead the r.v. exp(tX), where t is a real parameter to be used for optimizing the results.
- E[exp(tX)] = E[exp(t(X_1+...+X_k))] = E[Π_i exp(tX_i)]
- But, form the independence it follows that the expectation of the product is the product of the expectations:

$$E[\exp(tX)] = \prod_i E[\exp(tX_i)]$$

- So for t>0:

$$\Pr[X \geq a] = \Pr[\exp(tX) \geq \exp(ta)] \leq \frac{\prod_i E[\exp(tX_i)]}{\exp(ta)}$$

4

Using t<0:

$$\Pr[X \leq a] = \Pr[\exp(tX) \geq \exp(ta)] \leq \frac{\prod_i E[\exp(tX_i)]}{\exp(ta)}$$

To use Chernoff technique we need:

1) to be able to evaluate (upper bound) E[exp(tX_i)] for the class to which X_i belongs.

2) to choose t so minimize the bound.

5

Example (1)

- Given n independently identically distributed r.v. X_1,..,X_n that assume values +1, -1 with equal probability. Let X = $\Sigma_i X_i$. Clearly E(X)=0.
- Theorem: For any ϵ > 0:

$$\Pr[X \geq \varepsilon n] \leq \exp(-\frac{\varepsilon^2 n}{2})$$

- Proof: $E[\exp(tX_i)] = \frac{1}{2}\exp(t) + \frac{1}{2}\exp(-t) = \cosh(t) \leq \exp(\frac{t^2}{2})$
- the last inequality follows by comparing Taylor series.
- Applying Chernoff with a= ϵ n we get:

$$\Pr[X > \varepsilon n] \leq \exp(n(\frac{t^2}{2} - t\varepsilon))$$

- choose t = ϵ.

6

Example (2): binomial distribution p=1/2

Let $X = \Sigma_i X_i$, i=1..n, where $X_i = 0$ with probability 1/2 and 1 with probability 1/2.
We set up $Z_i = 2X_i - 1$, which has therefore values +1 and -1 with equal probability. We can apply the previous bound to $\Sigma_i Z_i$. Note that E[X]=n/2:

Theorem: for $\epsilon>0$: $\Pr[X - \frac{n}{2} \geq \varepsilon \frac{n}{2}] \leq \exp(-\frac{n\varepsilon^2}{2})$

In the case of a general Bernoulli distributionof parameter p, a more complex argument leads to a similar bound.

7

Example (3): geometric distribution

Suppose we have at our disposal a coin of bias p, we denote by X the number of trials needed to obtain a head. This r.v. X is distributed according to the geometric distribution.
More generally the number of trials needed to obtain n heads, T_n is called the negative binomial distribution. T_n can be expressed as sum of independent geometric distributions: $T_n = X_1+X_2+,...,+X_n$.
Thus we can use Chernoff.
$\mathrm{Prob}(X_i = x) = p(1-p)^{x-1}$, for $x \geq 1$. Thus

$$E[X_i] = \sum_{x\geq 1} xp(1-p)^{x-1} = \frac{1}{p}$$

So $E[T_n] = n/p$. We shall bound tails of T_n for p=1/2.

8

When p=1/2, $E[T_n]= 2n$.

Theorem: for $\epsilon\geq 3$: $\Pr[T_n \geq (2+\varepsilon)n] \leq \exp(-\frac{\varepsilon n}{4})$

Proof: let 0 < t < ln 2. We have

$$E[\exp(tX_i)] = \sum_{x\geq 1} \frac{\exp(tx)}{2^x} = \frac{\exp(t)}{2-\exp(t)}$$

We apply Chernoff with $a=(2+\epsilon)n$ obtaining:

$$\Pr[T_n \geq (2+\varepsilon)n] \leq \frac{1}{\exp(t(2+\varepsilon)n)}\left(\frac{\exp(t)}{2-\exp(t)}\right)^n = \left(\frac{\exp(-t(1+\varepsilon))}{2-\exp(t)}\right)^n$$

Set: $\exp(t) = 1+\frac{\varepsilon}{2+\varepsilon}, \quad 2-\exp(t) = \frac{1}{1+\varepsilon/2}, \quad \exp(-t) = 1-\frac{\varepsilon}{2+2\varepsilon}$

We get $\Pr[T_n \geq (2+\varepsilon)n] \leq \left((1+\varepsilon/2)\left(1-\frac{\varepsilon}{2+2\varepsilon}\right)^{1+\varepsilon}\right)^n$

9

Using the inequality $1-x \leq \exp(-x)$, which follows from the rule of alternating decreasing series, we bound:

$$\left(1-\frac{\varepsilon}{2+2\varepsilon}\right)^{1+\varepsilon} \leq \exp\left(-\frac{\varepsilon}{2+2\varepsilon}(1+\varepsilon)\right) = \exp(-\frac{\varepsilon}{2})$$

Also by direct calculation, for $\epsilon\geq 3$, $1+\varepsilon/2 \leq \exp(\varepsilon/4)$

And the bound of the Theorem follows.

10

Configuration Spaces

Given a set S of geometric objects (points, segments, halfplanes), build a derived structure H(S) [and G(S)] (decompostion of line into adjacent intervals, convex hull, intersection of halfsplanes, trapezoidal decomposition).

Incremental randomized method.
Build sequence $H(S_0), H(S_1),\ldots,H(S_n)$. Where $S_0=\emptyset$ and $S_{i+1}= S_i \cup \{p_i\}$ where $p_i \in S \setminus S_i$ chosen uniformely at random. Account for structural change and for conflict change.

11

Next to introduce the theory of random samples it is conveneint to embed this appraoch into an even more general combinatorial structure which is rich enough for proving theorems on the size of the conflict lists. Certain basic terms are left undefined on purpose.

Let Σ be a set of "configurations". Let S be a set of objects.
To Σ and S we associate two functions:
$T: \Sigma \to 2^S$, and $L: \Sigma \to 2^S$, with $T(\sigma) \cap L(\sigma)=\emptyset$.

$T(\sigma)$ intuitively is the set of triggers, that is elements of S that "define" σ.
$L(\sigma)$ intuitively us the set of stoppers, that is elements of S in "conflict" with σ.

12

Let $|T(\sigma)|$ be the degree of σ, and $|L(\sigma)|$ be the level of σ (also the conflict list size of σ. We partition the set Σ according to the value of the level.

$$\Pi_i(S) = \{\sigma \in \Sigma : |L(\sigma)| = i\}$$

Where $\Pi_0(S)=H(S)$ is called the set of active configurations. This is the set of configurations without conflicts, which is what we called H(S). So we find the old notion within the new terminology.

Definition. Let Σ be a configuration space. The *reduced configuration space* $<\Sigma>$ is the partition of Σ into equivalence classes with repect to the image of the T and L functions.

13

Also we denote with:

$$\Pi(S) = \bigcup_{i=0,\ldots,|S|} \Pi_i(S) = \Sigma$$

Consider a subset $R \subseteq S$ and we define the *restricted configuration space*:

$$\Pi(R) = \{\sigma \in \Sigma : T(\sigma) \subseteq R\}$$

And the *restricted conflict lists*:

$$L_R(\sigma) = L(\sigma) \cap R$$

Define the following subdivision of $\Pi(R)$ by levels :

$$\Pi_i(R) = \{\sigma \in \Sigma : |L_R(\sigma)| = i, T(\sigma) \subseteq R\}$$

14

Bounded *degree* configuration spaces:

A Configuration space Σ such that for a constant d, $|T(\sigma)| \leq d$, for all $\sigma \in \Sigma$.

Bounded *valence* configuration spaces:

A Configuration space Σ such that for a constant c, no more than c configurations $\sigma \in \Sigma$ have the same trigger and the same stopper sets.

Bounded *dimension* configuration spaces:

A Configuration space Σ such that for a constant t,

$$\max_{R \subseteq S, |R|=r} |< \Pi(R) >| = O(r^t)$$

15

With the new notation the structural change in phase i of an incremental algorithm is:

$$Str\cos t_i = |\Pi_0(S_i) \oplus \Pi_0(S_{i+1})|$$

And the conflict list change is:

$$Conf\ \cos t_i = \sum_{\sigma \in \Pi_0(S_i) \oplus \Pi_0(S_{i+1})} |L(\sigma)|$$

The *Total Stuctural Change Cost* and the *Total Conflict Lisit Cost* are the sums over all n phases.

16

Fundamental thorem of incremental constructions

Let Σ be a configuration space of bounded degree d, then an incremental randomized construction algorithm for $\Pi_0(S)$ has Expected Total Structural Change bounded by:

$$\sum_{i=1}^{n} \left(\frac{d}{i}\right) E[\Pi_0(S_i)]$$

The Expected Total Conflict List Change Cost is bounded by:

$$\sum_{i=1}^{n} \left(\frac{d^2(n-i)}{i^2}\right) E[\Pi_0(S_i)]$$

17